

## COLLISIONLESS SHOCKS IN A PARTIALLY IONIZED MEDIUM: III. EFFICIENT COSMIC RAY ACCELERATION

G. MORLINO<sup>1</sup>, P. BLASI<sup>1,2</sup>, R. BANDIERA<sup>1</sup>, E. AMATO<sup>1</sup>, D. CAPRIOLI<sup>3</sup>

<sup>1</sup>INAF-Osservatorio Astrofisico di Arcetri, Largo E. Fermi, 5, 50125 Firenze, Italy

<sup>2</sup>INFN/Laboratori Nazionali del Gran Sasso, Assergi, Italy and

<sup>3</sup>Department of Astrophysical Sciences, Peyton Hall, Princeton University, Princeton, NJ 08544, USA

*Draft version November 28, 2012*

### ABSTRACT

In this paper we present the first formulation of the theory of non-linear particle acceleration in collisionless shocks in the presence of neutral hydrogen in the acceleration region. The dynamical reaction of the accelerated particles, the magnetic field amplification and the magnetic dynamical effects on the shock are also included. The main new aspect consists however in accounting for charge exchange and ionization of neutral hydrogen, which profoundly change the structure of the shock, as discussed in our previous work. This important dynamical effect of neutrals is mainly associated to the so-called neutral return flux, namely the return of hot neutrals from the downstream region to the upstream, where they deposit energy and momentum through charge exchange and ionization. We also present the self-consistent calculation of Balmer line emission from the shock region and discuss how to use measurements of the anomalous width of the different components of the Balmer line to infer the cosmic ray acceleration efficiency in supernova remnants showing Balmer emission: the broad Balmer line, which is due to charge exchange of hydrogen atoms with hot ions downstream of the shock, is shown to become narrower as a result of the energy drainage into cosmic rays, while the narrow Balmer line, due to charge exchange in the cosmic-ray-induced precursor, is shown to become broader. In addition to these two well-known components, the neutral return flux leads to the formation of a third component with intermediate width: this too contains information on ongoing processes at the shock.

*Subject headings:* acceleration of particles – Cosmic Rays – Balmer emission – ISM: supernova remnants

### 1. INTRODUCTION

In the context of the so-called Supernova Remnant (SNR) paradigm for the origin of Galactic Cosmic Rays (CRs), the bulk of CRs are accelerated at SNR shocks through diffusive shock acceleration (DSA). Under this assumption, in order to explain the observed fluxes at Earth and, at the same time, the B/C ratio (sensitive to the grammage traversed by CRs during propagation) one is forced to require an acceleration efficiency of 5–15% per SNR, depending on details of the propagation in the Galaxy (e.g., Blasi & Amato 2012a,b). In these conditions, the non-linear effects induced by the dynamical reaction of accelerated particles on the SNR shock cannot be neglected and reflect on the shape of the spectrum (e.g., Malkov & Drury 2001, for a review) and on the temperature of the ionized plasma downstream of the shock (Decourchelle et al. 2000). Moreover, the streaming instability induced by accelerated particles ahead of the shock is thought to be responsible for the magnetic field amplification that is a crucial ingredient to reach maximum energies close to the knee (Blasi, Amato & Caprioli 2007) and to explain the thin non-thermal filaments observed in X-rays (e.g., Vink 2012). The non-linear theory of diffusive shock acceleration (NLDSA) including a phenomenological description of all these effects has been presented in a series of papers (Amato & Blasi 2005, 2006; Caprioli et al. 2009). Both test-particle DSA and NLDSA lead to predict spectra which are rather hard at high energies, with a power-law index in momentum  $\lesssim 4$ . This implies a steep dependence on energy of the Galactic diffusion coefficient (Berezhko & Völk 2007), if the CR spectrum at Earth is to be reproduced. Such energy dependence appears to be at odds with the observed anisotropy of Galactic CRs (Ptuskin et al. 2006; Blasi & Amato 2012b). Accounting for the velocity of scattering centers might lead to somewhat steeper injection spectra thereby mitigating the problem (Caprioli 2012).

In general, anything that affects the structure of the shock also leads to changes in the spectrum of accelerated particles. One important source of shock modification is represented by the presence of neutral atoms in the acceleration region: as discussed by Blasi et al. (2012, hereafter Paper I), the reactions of charge exchange (CE) and ionization produce important effects in the shock region. The most important of these effects is induced by the so-called *neutral return flux* (NRF): a hydrogen atom crossing a collisionless shock front can charge exchange with a hot ion downstream of the shock, thereby leading to the formation of a cold ion and a hot neutral. It may happen that the latter is produced with velocity in the direction of the shock; being neutral, it can recross the shock and suffer CE or ionization upstream, hence releasing energy and momentum there. This important phenomenon leads to heating of the upstream ionized plasma and to a decrease of the Mach number. Even in the context of test-particle DSA it is easy to show that this phenomenon leads to a steepening of the spectrum of particles accelerated in a SNR where appreciable neutral

hydrogen is present (Paper I). The NRF affects the upstream plasma on a scale of a few interaction lengths of CE.

The presence of neutral hydrogen also represents a unique diagnostic tool of the acceleration process, through the Balmer emission associated with CE to excited states: in a Balmer dominated shock (see Heng 2009, for a review) the Balmer line has a narrow and a broad component, the first associated with hydrogen atoms that did not suffer CE with higher temperature ions, and the second that reflects the temperature of shocked ions, due to CE reactions downstream (Chevalier & Raymond 1978; van Adelsberg et al. 1978). In the presence of CR acceleration the ions' temperature downstream is lower, and this should reflect in decreased Full-Width at Half-Maximum (FWHM) of the broad Balmer line. If the CR pressure is large enough, a CR-induced precursor is formed in front of the shock (Malkov & Drury 2001) so that CE can also take place upstream and eventually lead to a broadening of the narrow Balmer line. In addition, it was shown in Morlino et al. (2012, hereafter Paper II), that the shape of the Balmer line is also affected by the NRF discussed above through the formation of an intermediate component of the line.

Remarkably, anomalous widths of both broad and narrow lines have been observed in Balmer-dominated shocks. For instance, Helder et al. (2009) combined proper-motion measurements of the shock and broad  $H\alpha$  line width for the remnant RCW 86 to demonstrate that the inferred low temperature behind the shock may be interpreted as a signal that a sizable fraction of the energy is being channeled into CRs. A similar conclusion has been reached by Helder et al. (2010) by analyzing the SNR 0509-67.5 in the LMC, even if in this case the shock velocity has been estimated using an evolution model with information derived from the line width of elements in the shocked ejecta (which can be used to infer the speed of reverse shock).

Anomalous width of narrow Balmer lines have been also observed in several SNRs (see, e.g. Sollerman et al. 2003). The width of such lines is in the  $30\text{--}50\text{ km s}^{-1}$  range, implying a pre-shock temperature around  $25,000\text{--}50,000\text{ K}$ . If this were the ISM equilibrium temperature there would be no atomic hydrogen, implying that the pre-shock hydrogen is heated by some form of shock precursor in a region that is sufficiently thin so as to make collisional ionization equilibrium before the shock unfeasible. The CR precursor is the most probable candidate to explain such a broadening of the narrow line width, though other mechanisms have also been proposed (see Heng 2009, for a review).

A previous attempt to include neutral particles in the shock acceleration theory has been done by Morlino et al. (2010) using a fluid approach. Nevertheless, as discussed in Paper I, neutrals cannot be treated as a fluid because the length scales involved are much smaller than the equilibration scale. More recently Raymond et al. (2011) tried to describe the interaction of neutrals with the CR precursor using a phenomenological approach where both the CR spectrum and spatial distribution are assumed rather than calculated.

An accurate description of the CR acceleration process in the presence of neutrals can only be achieved by formulating a NLDSA theory where the dynamical action of neutrals is taken into account. This is a needed preliminary step, if one wants to make use of measurements of Balmer line emission from SNRs to quantitatively assess their efficiency as particle accelerators. In this paper we present the first treatment of NLDSA in partially ionized media and discuss how to use the theory to calculate the shape of the Balmer line and infer information on particle acceleration at the shock.

The paper is structured as follows: in Section 2 we describe the calculations used to describe neutrals, ions and the acceleration process and we explain in detail how to assemble the parts together to describe NLDSA in the presence of neutrals. The formalism previously introduced in Paper I — as far as neutrals are concerned — and in (Amato & Blasi 2005, 2006; Caprioli et al. 2009) — as far as NLDSA is concerned — will be heavily used. In Section 3 we illustrate our results in terms of shock modification, spectrum of accelerated particles and shape of the Balmer line. We summarize and conclude in Section 4.

## 2. CALCULATIONS

We consider a plane-parallel shock wave propagating in a partially ionized proton-electron plasma with velocity  $V_{\text{sh}}$  along the  $z$  direction. The fraction of neutral hydrogen is fixed at upstream infinity where ions and neutrals are assumed to be in thermal equilibrium with each other. The shock structure is determined by the interaction of CRs and neutrals with the background plasma. Both CRs and neutrals profoundly change the shock structure, especially upstream where both create a precursor: the CR-induced precursor reflects the diffusion properties of accelerated particles and has a typical spatial scale of the order of the diffusion length of the highest energy particles. The neutral-induced precursor develops on a spatial scale comparable with a few interaction lengths of the dominant process between CE and ionization. The downstream region is also affected by the presence of both CRs and neutrals and the velocity gradients that arise from ionization have a direct influence on the spectrum of accelerated particles.

A self consistent description of the overall system requires to consider four mutually interacting components: thermal particles (protons and electrons), neutrals (hydrogen), accelerated protons (CRs) and turbulent magnetic field. We neglect the presence of helium and heavier chemical elements. In the following subsections we write down the basic equations used to describe each component, including the interaction terms. We are looking for stationary solutions, hence all equations are written as time-independent. The interaction terms make the system highly non linear and the solution is found using a iterative scheme similar to the one we introduced in some of our previous work (Blasi 2002; Amato & Blasi 2005, Paper I). The details on how to generalize such scheme to the case in which neutrals are present are discussed in § 2.5.

### 2.1. Neutral distribution

The behavior of neutrals has been extensively discussed in Paper I. Neutral hydrogen interacts with protons through CE and ionization and with electrons through ionization only. The hydrogen distribution function,  $f_N(\vec{v}, z)$ , can be

described using the stationary Boltzmann equation

$$v_z \frac{\partial f_N(\vec{v}, z)}{\partial z} = \beta_N f_i(\vec{v}, z) - [\beta_i + \beta_e] f_N(\vec{v}, z), \quad (1)$$

where  $z$  is the distance from the shock (which is located at the origin),  $v_z$  is the velocity component along the  $z$  axis and the electron and proton distribution functions,  $f_i(\vec{v}, z)$  and  $f_e(\vec{v}, z)$ , are assumed to be Maxwellian at each position. The collision terms,  $\beta_k f_l$ , describe the interaction (due to CE and/or ionization) between the species  $k$  and  $l$ . The interaction rate  $\beta_k$  is formally written as

$$\beta_k(\vec{v}, z) = \int d^3 w v_{rel} \sigma(\vec{v}_{rel}) f_k(\vec{w}, z), \quad (2)$$

where  $v_{rel} = |\vec{v} - \vec{w}|$  and  $\sigma$  is the cross section for the relevant interaction process. More precisely,  $\beta_N$  is the rate of CE of an ion that becomes a neutral,  $\beta_i$  is the rate of CE plus ionization of a neutral due to collisions with protons, while  $\beta_e$  is the ionization rate of neutrals due to collisions with electrons. A full description of the cross sections used in the calculations can be found in Paper II.

A novel method to solve Eq. (1) has been developed in Paper I, where we showed how to calculate  $f_N$  starting from the distribution of protons and electrons. The method is based on decomposing the distribution function  $f_N$  in the sum of the neutrals that have suffered 0, 1, 2, ...,  $k$  processes of charge exchange. Each distribution function is named  $f_N^{(k)}$ , and clearly  $f_N = \sum_{k=0}^{\infty} f_N^{(k)}$ . Using this formalism, Eq. (1) can be rewritten as a set of  $k + 1$  equations, one for each component:

$$v_z \partial_z f_N^{(0)} = -[\beta_i + \beta_e] f_N^{(0)} \quad \text{for } k = 0, \quad \text{and} \quad (3)$$

$$v_z \partial_z f_N^{(k)} = \beta_N^{(k-1)} f_i - [\beta_i + \beta_e] f_N^{(k)} \quad \text{for } k > 0. \quad (4)$$

Each one of these equations is linear, and the solution for  $f_N^{(k)}$  can be obtained from  $f_N^{(k-1)}$ . The boundary conditions are imposed at upstream infinity where ions and neutrals are assumed to start with the same bulk velocity and temperature, therefore charge exchange occurs at equilibrium and the distributions do not change. As a consequence, the boundary conditions are  $f_N^{(0)}(\vec{v}, -\infty) = (n_{N,0}/n_{i,0}) f_i(\vec{v}, -\infty)$  and  $f_N^{(k>0)}(\vec{v}, -\infty) = 0$ . A good approximation to the total distribution function of neutrals is obtained when a sufficient number of  $f_N^{(k)}$  is taken into account. This number is determined by the physical scale of the problem. A rough estimate can be obtained comparing the lengthscale for CE with the maximum between the ionization lengthscale and the precursor length. In general, for reasonable values of the ambient parameters, the longest scale, and the one relevant for comparison, is that of the CR-induced precursor.

## 2.2. CR distribution

The isotropic distribution function of CRs satisfies the following transport equation in the reference frame of the shock:

$$\frac{\partial}{\partial z} \left[ D(z, p) \frac{\partial f_c}{\partial z} \right] - u \frac{\partial f_c}{\partial z} + \frac{1}{3} \frac{du}{dz} p \frac{\partial f_c}{\partial p} + Q(z, p) = 0. \quad (5)$$

The  $z$ -axis is oriented from upstream infinity ( $z = -\infty$ ) to downstream infinity ( $z = +\infty$ ) with the shock located at  $z = 0$ . We assume that the injection occurs only at the shock position and is monoenergetic at  $p = p_{inj}$ , hence the injection term can be written as  $Q(z, p) = Q_0(p) \delta(z)$  where  $Q_0(p) = (\eta_{inj} n_1 / 4\pi p_{inj}^2) \delta(p - p_{inj})$ . Here  $n_1$  is the number density of ions immediately upstream of the subshock, and  $\eta_{inj}$  is the fraction of particles that is going to take part in the acceleration process. Following Blasi et al. (2005),  $\eta_{inj}$  can be related to the subshock compression factor as:

$$\eta_{inj} = 4 / (3\sqrt{\pi}) (R_{sub} - 1) \xi_{inj}^3 e^{-\xi_{inj}^2}. \quad (6)$$

Here  $\xi_{inj}$  is defined by the relation  $p_{inj} = \xi_{inj} p_{th,2}$ , where  $p_{th,2}$  is the momentum of the thermal particles downstream.  $\xi_{inj}$  parametrizes the poorly known microphysics of the injection process and is taken as a free parameter with a typical value between 2 and 4.

The diffusion properties of particles are described by the diffusion coefficient  $D(z, p)$ . We assume Bohm diffusion in the local amplified magnetic field:

$$D(z, p) = \frac{1}{3} c r_L [\delta B(z)], \quad (7)$$

where  $r_L(\delta B) = pc/[e\delta B(z)]$  is the Larmor radius in the amplified magnetic field. The calculation of  $\delta B$  is described in § 2.3.

The solution of Eq. (5) can be easily obtained by using standard techniques. Here we generalize the approach of Blasi (2002) and Amato & Blasi (2005) in order to relax the assumption of null gradient of the distribution function downstream (homogeneity), which does not apply to the general case in which CRs and neutrals are present.

The solution of Eq. (5) in the upstream is the same found by Amato & Blasi (2005). Such a solution can be well approximated by the following expression (see also Caprioli et al. 2010):

$$f_{up}(z, p) = f_0(p) \exp \left[ - \int_z^0 dz' \frac{u(z')}{D(z', p)} \right], \quad (8)$$

where  $f_0(p) = f(0, p)$  is the distribution function at the shock position. Through the same approach used by Amato & Blasi (2005) we can express the solution in the downstream in the following implicit form:

$$f_{down}(z, p) = f_0(p) + \frac{1}{3} \int_0^z \frac{dz'}{D} \int_{z'}^\infty dz'' \frac{du}{dz''} p \frac{\partial f}{\partial p} \exp \left[ - \int_{z'}^{z''} \frac{u}{D} dy \right]. \quad (9)$$

We found that an excellent approximation to the expression in Eq. (9) is given by:

$$f_{down}(z, p) = f_0(p) \exp \left\{ - \frac{q_0(p)}{3} \int_0^z \frac{dz'}{D} \int_{z'}^\infty dz'' \frac{du}{dz''} \exp \left[ - \int_{z'}^{z''} \frac{u}{D} dz'' \right] \right\}, \quad (10)$$

where  $q_0(p) = -d \ln f_0(p) / d \ln p$  is the slope of the distribution function at the shock position. Both Equations (9) and (10) give  $f(z, p) = f_0(p)$  when the downstream speed  $u(z)$  is constant.

The distribution function at the shock is obtained, as usual, by integrating Eq. (5) across the subshock:

$$\frac{u_1 - u_2}{3} p \frac{\partial f_0}{\partial p} = \left[ D \frac{\partial f}{\partial z} \right]_1 - \left[ D \frac{\partial f}{\partial z} \right]_2 + Q_0(p). \quad (11)$$

Here and in the rest of the paper, labels “1” and “2” refer to quantities immediately upstream and immediately downstream of the subshock, respectively. The value of  $[D \partial f / \partial z]_1$  can be obtained by integrating Eq. (5) between upstream infinity ( $z = -\infty$ ) and position 1 (Blasi 2002). In the standard stationary solution the downstream distribution is homogeneous, hence  $[D \partial f / \partial z]_2 \equiv 0$ . This condition does not hold in our case because of ionization of neutrals downstream. The general expression can be derived from Eq. (9) and reads

$$\left[ D \frac{\partial f}{\partial z} \right]_2 = -f_0(p) \frac{q_0(p)}{3} \int_0^\infty dz \frac{du}{dz'} \exp \left[ - \int_0^{z'} \frac{u}{D} dz'' \right]. \quad (12)$$

Substituting the last expression in Eq. (11) we obtain for  $f_0(p)$ :

$$f_0(p) = \frac{\eta n_0}{4\pi p_{inj}^3} \frac{3R_{tot}}{R_{tot} U_p(p) - 1} \exp \left\{ - \int_{p_{inj}}^p \frac{dp'}{p'} \frac{3R_{tot}}{R_{tot} U_p(p') - 1} \left[ U_p(p') + \frac{q_0(p')}{3} K(p') \right] \right\}, \quad (13)$$

where we defined  $U_p(p) = u_p/V_{sh}$  and  $u_p(p)$  can be interpreted as an estimate of the mean plasma velocity that a particle with momentum  $p$  experiences in the precursor region (see Equation (8) in Amato & Blasi 2005). Eq. (13) differs from the standard solution only for the term including  $K(p)$ , where

$$K(p) = \frac{1}{V_{sh}} \int_0^\infty dz \frac{du}{dz} \exp \left[ - \int_0^z \frac{u}{D} dz' \right]. \quad (14)$$

The mean plasma speed experienced in the downstream by particles that are able to return to the subshock is  $u_2 + V_{sh} K(p)$ . When the downstream plasma has a constant velocity,  $K(p) \rightarrow 0$  and Eq. (13) reduces to the standard solution (see Equation (7) in Amato & Blasi 2005).

### 2.3. Transport equation for waves

Following Amato & Blasi (2006) we describe the scattering of accelerated particles in the acceleration region as due to the waves that the particles generate through resonant streaming instability. These waves are also damped due to several processes. In particular, when the plasma is not fully ionized, the presence of neutrals can damp Alfvén waves via ion-neutral damping.

The equation for transport of waves can be written as:

$$\frac{\partial F_w(k, z)}{\partial z} = u(z) \frac{\partial P_w(k, z)}{\partial z} + P_w(k, z) [\sigma_{CR}(k, z) - \Gamma_{TH}(k, z)], \quad (15)$$

where  $F_w$  and  $P_w$  are, respectively, the energy flux and the pressure per unit logarithmic bandwidth of waves with wavenumber  $k$ .  $\sigma$  is the growth rate of magnetic turbulence, while  $\Gamma_{TH}$  is the damping rate. For resonant wave amplification the growth rate of Alfvén waves is (Bell 1978):

$$\sigma_{CR}(k, x) = \frac{4\pi}{3} \frac{v_A(x)}{P_w(k, x)} \left[ p^4 v(p) \frac{\partial f_{CR}}{\partial x} \right]_{p=\bar{p}(k)}, \quad (16)$$

where  $p = \bar{p}(k) = eB/km_pc$  is the resonant momentum. The damping of the waves is mainly due to non-linear Landau damping and ion-neutral damping. For the sake of simplicity here we adopt a phenomenological approach in which the damping results in a generic Alfvén heating at a rate  $\dot{\Gamma}_{\text{TH}} = \eta_{\text{TH}}\sigma_{\text{CR}}$ . This expression assumes that a fraction  $\eta_{\text{TH}}$  of the power in amplified waves is locally damped and results in heating of the background plasma (Berezhko & Ellison 1999).

Following Caprioli et al. (2009), Eq. (15) can be solved by integrating in  $k$ -space and using the relation between energy density and pressure of Alfvén waves, namely  $F_w = 3P_wu$ , which holds upstream when  $v_A \ll V_{\text{sh}}$ . This leads to the following equation:

$$2U(z) \frac{d\tilde{P}_w(z)}{dz} = V_A(z) [1 - \eta_{\text{TH}}] \frac{d\tilde{P}_{\text{CR}}(z)}{dz} - 3\tilde{P}_w(z) \frac{dU(z)}{dz}, \quad (17)$$

where we used the normalized quantities  $U = u/V_{\text{sh}}$ ,  $V_A = v_a/V_{\text{sh}}$ ,  $\tilde{P} = P/(\rho_0 V_{\text{sh}}^2)$ . Eq. (17) can be used to derive  $P_w$  in the upstream region.

The jump conditions for the magnetic field are calculated following Caprioli et al. (2009):

$$P_{w,2} = R_{\text{sub}} P_{w,1}, \quad (18)$$

$$F_{w,2} - F_{w,1} = 2(R_{\text{sub}} - 1)u_{i,1}P_{w,1}. \quad (19)$$

In the downstream section we solve the same transport Eq. (15) but we assume the relation between energy flux and pressure that derives from Eqs. (18) and (19), namely  $F_w = (1 + 2R_{\text{sub}})uP_w$ . Moreover, in the downstream region we assume for simplicity that damping occurs on scales much larger than the typical scale length for CE and ionization (this assumption is not strictly needed, it simply makes the results of the calculations easier to interpret).

#### 2.4. Dynamics of the ionized gas

The dynamics of the background plasma is affected by the presence of accelerated particles and by CE and ionization of neutrals. Protons and electrons in the plasma are assumed to share the same local density,  $\rho_i(z) = \rho_e(z)$ , but not necessarily the same temperature, i.e.,  $T_i(z)$  may be different from  $T_e(z)$ . The equations describing the conservation of mass, momentum and energy taking into account the interactions of the plasma fluid with CRs are:

$$\frac{\partial}{\partial z} [\rho_i u_i + \mu_N] = 0, \quad (20)$$

$$\frac{\partial}{\partial z} [\rho_i u_i^2 + P_g + P_c + P_w + P_N] = 0, \quad (21)$$

$$\frac{\partial}{\partial z} \left[ \frac{1}{2} \rho_i u_i^3 + \frac{\gamma_g P_g u_i}{\gamma_g - 1} + F_w + F_N \right] = -u_i \frac{\partial P_c}{\partial z} + \Gamma P_w. \quad (22)$$

Here  $\mu_N = m_H \int d^3v v_{\parallel} f_N$ ,  $P_N = m_H \int d^3v v_{\parallel}^2 f_N$  and  $F_N = m_H/2 \int d^3v (v_{\parallel}^2 + v_{\perp}^2) f_N$  are respectively the fluxes of mass, momentum and energy of neutrals along the  $z$  direction (labelled as  $\parallel$ ). They can be easily computed once the neutral distribution function is known.  $P_w$  and  $F_w$  are the pressure and energy flux of waves as defined in section § 2.3, while  $P_c$  is the CR pressure computed from the CR distribution function:

$$P_c(z) = \frac{4\pi}{3} \int dp p^3 v(p) f_c(z, p). \quad (23)$$

The dynamical role of electrons in the conservation equations is usually neglected due to their small mass. However, collective plasma processes could contribute to equilibrate electron and proton temperatures, at least partially. If the equilibration occurs in a very efficient manner, the electron pressure cannot be neglected and the total gas pressure needs to include both the proton and electron contributions, namely  $P_g = P_i + P_e = P_i(1 + \beta)$ , where  $\beta(z) \equiv T_e/T_i$  is the electron to proton temperature ratio and is taken here as a free parameter. We note that the electron-ion equilibration level can be different between upstream and downstream because the plasma conditions in these two regions are different. Hence we distinguish between upstream and downstream using two separate parameters,  $\beta_{\text{up}}$  and  $\beta_{\text{down}}$ , respectively. On the right side of Eq. (22) we include the heating produced by the compression of CR acting on the plasma plus the heating due to the damping of magnetic waves.

The computation strategy we adopt is as follows: we first solve Eqs. (20)-(22) for the upstream dynamics starting from the boundary conditions at upstream infinity, i.e.  $\rho_i(-\infty) = \rho_{i,0}$  and  $u_i(-\infty) = V_{\text{sh}}$ . In order to determine the conditions immediately downstream of the subshock we use the jump conditions derived from the conservation equations:

$$[\rho_i u_i + \mu_N]_1^2 = 0, \quad (24)$$

$$[\rho_i u_i^2 + P_g + P_w]_1^2 = 0, \quad (25)$$

$$\left[ \frac{1}{2} \rho_i u_i^3 + \frac{\gamma_g P_g u_i}{\gamma_g - 1} + F_w \right]_1^2 = 0, \quad (26)$$

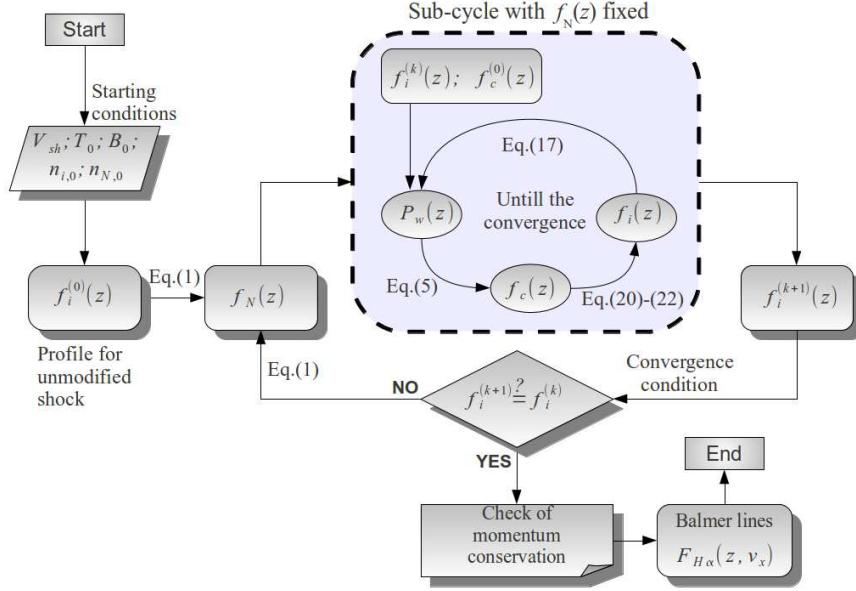


FIG. 1.— Flow chart of the iterative method used to find the solution.

where the brackets indicate the difference between quantities upstream and downstream of the shock ( $[X]_1^2 = X_2 - X_1$ ). The terms describing neutrals and CRs are absent because of the continuity of  $f_N$  and  $f_c$  across the subshock. Finally we solve the conservation equations, Eqs. (20)-(22) in the downstream region, using the values of the quantities at position 2 as boundary conditions.

### 2.5. Iterative method for the solution of the non-linear problem

In order to solve the set of non-linear equations involving neutrals, ions, CRs and magnetic field, we adopt an iterative method that we illustrate in Fig. (1) in the form of a flow diagram. The input quantities are the values of the shock velocity and all environmental quantities at upstream infinity, where the distribution function of neutrals is assumed to be Maxwellian at the same temperature as that of ions. The method proceeds through two nested cycles of iteration: in the first cycle, we start with a distribution function of ions that corresponds to an unperturbed distribution that one would obtain in the absence of both neutrals and CRs and we calculate the neutral distribution function  $f_N(z)$  from Eq. (1). With the given  $f_N(z)$  we iterate to obtain the ion distribution, the CR distribution and the wave energy density. The outer loop of iteration allows us to update the neutral distribution function using the newly available ion and CR distribution functions. The iteration ends when the predefined accuracy is achieved. At this point the distribution functions are used to calculate the structure of the Balmer line, following the method illustrated in Paper II.

The typical computation time required to complete the calculations for a given set of environment parameters is of the order of one hour on a standard laptop. The most time-consuming part is the solution of Eq. (1) which increases linearly with the number of partial functions,  $f_N^{(k)}$ , required for a correct description of the neutral distribution function. In particular the number of  $f_N^{(k)}$  increases linearly with the ratio between the precursor length and the CE interaction length.

## 3. RESULTS

The formalism presented above represents the first theory of particle acceleration at collisionless shocks in the presence of neutral atoms and allows us to calculate the shock structure, the spectrum of accelerated particles and the Balmer emission from the shock region. In the subsections below we describe the main results concerning these three topics.

### 3.1. Shock structure and spectra of accelerated particles

We adopt a test case consisting of a shock moving at  $V_{sh} = 4,000 \text{ km s}^{-1}$  in a medium with temperature  $T_0 = 10^4 \text{ K}$ , magnetic field  $B_0 = 10 \mu\text{G}$ , and densities  $\rho_{i,0} = \rho_{N,0} = 0.05 \text{ cm}^{-3}$  (the neutral fraction is 50%). CRs are injected with an efficiency of  $\sim 20\%$  (in fact we fix the injection parameter  $\xi = 3.7$ ) and the efficiency of Turbulent Heating (TH hereafter) is fixed as  $\eta_{TH} = 0.5$  (see Eq. 17). The shock structure is illustrated in Fig. 2: the top panel shows the pressure of gas (continuous solid line), magnetic field (dotted line) and CRs (dashed line) as functions of  $z$ , plotted

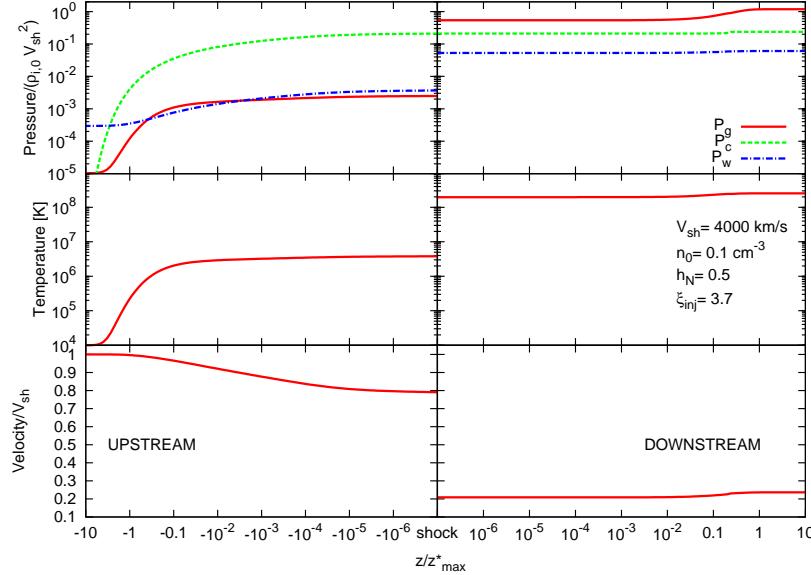


FIG. 2.— Example of spatial profiles of relevant quantities for a shock propagating in partially neutral plasma. Top panel: pressure of ions, CRs and magnetic field. Middle panel: ion temperature. Bottom panel: ion velocity. Quantities are shown in the shock frame with the subshock located at the origin. The  $z$  axis is shown in logarithmic scale both upstream and downstream and  $z_{\max}^* = D(p_{\max})/V_{sh}$  is the diffusion length of particles with the maximum momentum. These profiles are computed for:  $V_{sh} = 4,000 \text{ km s}^{-1}$ ,  $T_0 = 10^4 \text{ K}$ ,  $B_0 = 10 \mu\text{G}$ ,  $\rho_{i,0} = \rho_{N,0} = 0.05 \text{ cm}^{-3}$ . The injection parameter for CRs is  $\xi_{\text{inj}} = 3.7$ , which corresponds to an acceleration efficiency  $\sim 20\%$ .

here in logarithmic scale for both the upstream and downstream regions. All pressure components are normalized to the ram pressure  $\rho_{i,0} V_{sh}^2$  at upstream infinity. Clearly the CR pressure is continuous across the shock front. The self-generated magnetic field acquires a pressure that is of the same order of magnitude of the thermal pressure, which implies that its dynamical effect is non-negligible, as already discussed in Caprioli et al. (2008, 2009).

The background ions are substantially heated in the precursor due to TH, as it is shown by the increase in the gas temperature on a spatial scale comparable with  $z_{\max}^* \equiv z^*(p_{\max}) \equiv D(p_{\max})/V_{sh}$ , i.e., the diffusion length of particles at  $p_{\max}$  (mid panel of Fig. 2). The upstream profile of the fluid velocity, showed in the bottom panel of Fig. 2 as normalized to  $V_{sh}$ , is a distinctive signature of the CR-induced precursor.

Finally, one last thing to notice in this Figure is the increase of both the gas pressure and velocity at some distance from the shock in the downstream. This is a result of progressive ionization of fast cold neutrals, and indeed occurs on scales which are of the order of the ionization length. We will see that the situation is different, with the plasma velocity decreasing, rather than increasing, towards downstream infinity, when the shock velocity is lower and the NRF starts to play an important role (see Fig. 4 below). Spatial variations of the plasma velocity are important for determining the accelerated particle spectrum at the shock, whose slope at each given energy always reflects the mean compression ratio experienced by particles of that energy while moving away from the shock.

The spectrum of accelerated particles at different locations upstream and downstream of the shock is shown in Fig. 3 for a neutral fraction  $h_N = 0.5$ . The solid lines refer to the downstream section of the plasma, the thick dot-dashed line is the spectrum at the subshock location, the dashed lines are the spectra at different locations upstream of the shock.

One can see that, moving away from the shock in the upstream direction, the spectrum is more and more depleted of the lowest energy particles: at any position  $z$ , only particles energetic enough to have a diffusion length  $D(p)/V_{sh} \gtrsim z$  are present. The transport of low-energy particles in the downstream, is rather dominated by advection, and the spectrum is never truncated: at downstream infinity, the low-energy spectrum has the same slope as at the shock, but with a higher normalization. This is due to the compression that occurs around  $0.1 < z/z_{\max}^* < 1$ , where ionization takes place. At the highest energies, instead, diffusion dominates over advection and the CR spectrum far downstream is expected (and found) to be exactly the same as at the shock location. As a result, at large distances downstream of the shock the spectrum turns out to be slightly steeper than at the shock.

As already illustrated in Paper I in the test-particle limit, the effect of the NRF is to create an upstream precursor that steepens the spectra of accelerated particles. For this reason the spectra of accelerated particles are less concave than they would be in the presence of the CR reaction alone. This result is qualitatively general, though the strength of the effect is quantitatively dependent upon the specific environment that the shock propagates into, as discussed below.

In Fig. 4 we show the normalized plasma velocity as a function of the distance from the shock for different values of the neutral fraction (left panel), and the slope of the spectrum of accelerated particles as a function of momentum (right panel). All the calculations refer to a shock speed  $V_{sh} = 2,000 \text{ km s}^{-1}$ , a total gas density at upstream infinity  $n_0 = 0.1 \text{ cm}^{-3}$ , an unperturbed magnetic field strength  $B_0 = 10 \mu\text{G}$  and an injection parameter  $\xi_{\text{inj}} = 3.7$ . The solid

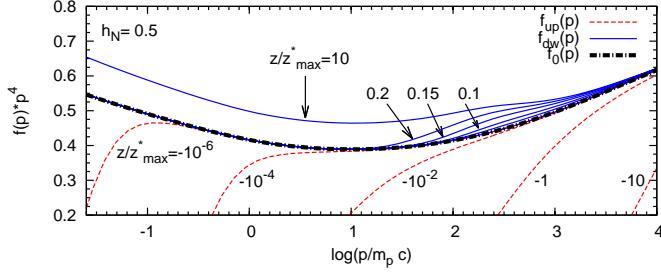


FIG. 3.— CR momentum spectrum multiplied by  $p^4$  for different distances from the sub-shock: the dash-dotted (black) line is the spectrum at the sub-shock position, the dashed (red) lines show the spectrum in the upstream and solid (blue) lines show the spectrum downstream. The position  $z$  is indicated for each curve in units of  $z^*_{\max}$ .

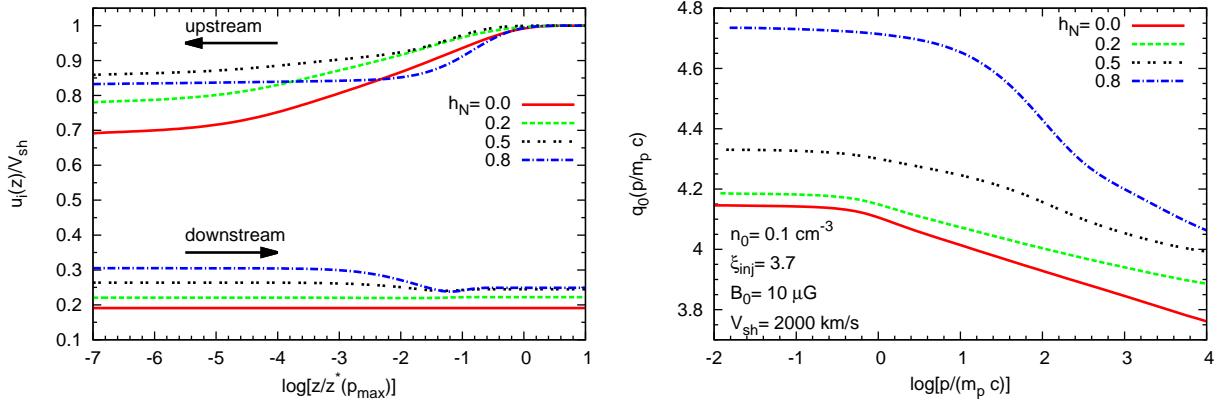


FIG. 4.— *Left panel*: Velocity profile of the upstream and downstream fluids normalized to the shock velocity, as a function of the distance from the shock. The shock is located on the left side and the (absolute) distance is in unit of  $z^*(p_{\max})$ . The profiles are shown for a given shock speed  $V_{sh} = 2,000 \text{ km s}^{-1}$  and for different neutral fractions  $h_N$  from 0 up to 80% as in the legend. Other parameters are  $\xi_{inj} = 3.7$ ,  $B_0 = 10 \mu\text{G}$  and  $n_0 = 0.1 \text{ cm}^{-3}$ . *Right panel*: Slope of the spectrum of accelerated particles for the same shock configurations as in the left panel.

lines refer to the case with fully ionized plasma ( $h_N = 0$ ). In this case the shock is appreciably modified by the CR pressure, the acceleration efficiency being  $\sim 30\%$ , and the spectrum is quite concave. The spectrum is steeper than  $p^{-4}$  only for energies below  $\sim 10$  GeV.

Due to the relatively-low shock velocity, the effects of CE and ionization are quite relevant; therefore, we expect the results to change appreciably for non vanishing values of  $h_N$ , due to the NRF. This is illustrated by the long-dashed, dashed and dotted lines in Fig. 4. For  $h_N = 0.2$  (long-dashed line) and  $h_N = 0.5$  (dashed line) the formation of the precursor upstream is dominated by the dynamical reaction of CRs on the plasma, although the heating of the upstream due to the NRF is very evident. In these two cases the acceleration efficiency is  $\sim 20\%$  (for  $h_N = 0.2$ ) and  $\sim 15\%$  (for  $h_N = 0.5$ ). The spectra in the same cases are correspondingly less modified by the CR pressure (less concave) and, on average, somewhat steeper: for  $h_N = 0.2$  the spectrum is steeper than  $p^{-4}$  for energies  $\lesssim 100$  GeV, while for  $h_N = 0.5$  it remains steeper than  $p^{-4}$  even at  $E \sim 10^4$  GeV.

The case with large neutral fraction,  $h_N = 0.8$  is rather extreme: the precursor becomes very short, as it becomes dominated by the NRF, and the acceleration efficiency is still  $\sim 15\%$ . The spectrum is quite steeper than  $p^{-4}$  at all energies, so that the bulk of the pressure is carried by relatively low energy particles. This implies that the CR pressure can slow down the incoming plasma only very close to the subshock (see left panel in Fig. 4).

As discussed in Paper I, the effects of the NRF are most important for shock speeds  $V_{sh} \leq 3,000 \text{ km s}^{-1}$ ; for larger shock speeds, the typical relative velocity between neutrals and ions is such that the cross section for CE is so small that atoms get ionized before suffering a CE reaction. As a consequence, the NRF is suppressed. It is useful to explore the dependence of the spectrum on the shock velocity in a case in which the neutral fraction is  $h_N = 0.5$ . In the left panel Fig. 5 we plot the normalized plasma velocity as a function of distance from the shock for shock velocities  $V_{sh} = 2,200 \text{ km s}^{-1}$  (solid line),  $V_{sh} = 3,000 \text{ km s}^{-1}$  (dashed line) and  $V_{sh} = 4,000 \text{ km s}^{-1}$  (dotted line). The other parameters are the same as in Fig. 4. The right panel shows the spectral slope for the same values of the shock velocity. For comparison, the thick lines refer to  $h_N = 0.5$ , while the thin lines of the same type refer to completely ionized gas ( $h_N = 0$ ).

The left panel qualitatively confirms our expectations based on the test-particle results: when increasing the shock speed from  $2,200 \text{ km s}^{-1}$  to  $4,000 \text{ km s}^{-1}$ , the effect of neutrals becomes smaller, namely the shock becomes increasingly

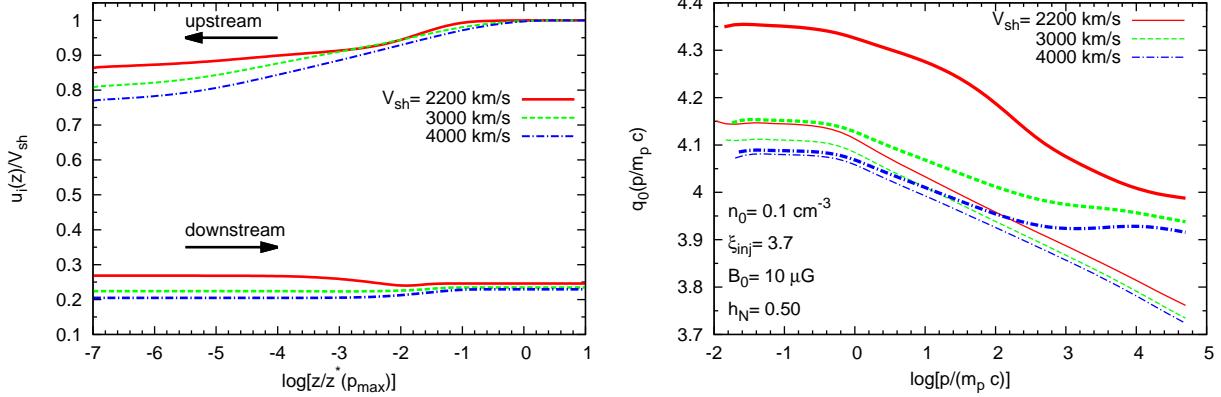


FIG. 5.— As in Fig. 4, but for a fixed neutral fraction  $h_N = 0.5$  and different shock velocities, as indicated. For comparison, in the right panel the thin lines show the slope in the absence of neutrals.

more modified by the reaction of CRs. This is even more evident in the right panel of Fig. 5: for  $V_{sh} = 2,200 \text{ km s}^{-1}$  a neutral fraction of 50% is sufficient to make the spectrum of accelerated particles steeper than  $p^{-4}$ , at all energies (thick solid line). With the same shock speed, but for a completely ionized plasma, the spectrum is instead harder than  $p^{-4}$  above  $\sim 10 \text{ GeV}$  ( $h_N = 0$ , thin solid line).

### 3.2. Balmer emission

The Balmer line is affected by the presence of CRs in at least two ways. First of all, the width of the broad Balmer line is expected to be narrower if efficient CR acceleration takes place. This is a straightforward consequence of energy conservation: energy is channelled into accelerated particles, therefore the plasma temperature downstream is lower. Moreover, the formation of a CR-induced precursor leads to a non-negligible heating upstream of the subshock, so that the narrow Balmer line may become broader. In this section we investigate these possibilities in a more quantitative way, by using the formalism introduced in Paper II. The different components of the line are identified as discussed in Paper II. The global shape of the Balmer line is fitted with three Gaussian components: a narrow line, an intermediate line (reflecting directly the effect of the NRF) and a broad line.

In Fig. 6 we show the temperature throughout the shock region for different values of the efficiency of TH,  $\eta_{TH}$ , for a shock propagating at  $4,000 \text{ km s}^{-1}$  in a medium with density  $0.1 \text{ cm}^{-3}$ . The case without CR acceleration is shown as a thick solid (black) line. In this case the precursor is solely due to the NRF and is limited to the region very close to the subshock.

When the acceleration process is turned on, several phenomena become visible. In the absence of TH ( $\eta_{TH} = 0$ , thin solid line), the upstream fluid is mildly heated by the adiabatic compression due to the CR-induced precursor only. Nevertheless, the temperature of the downstream plasma drops by almost a factor 2 as a consequence of the energy channelled into the accelerated particles. Increasing the value of  $\eta_{TH}$ , the heating in the precursor becomes more and more marked, and it extends over larger and larger distances. However, the downstream temperature is not appreciably affected by a change in the efficiency of TH.

The shape of the global Balmer line emission is plotted in the left panel of Fig. 7, while the right panel shows a zoom in the region of the narrow Balmer line. The curves refer to the same cases as in Fig. 6, labelled as indicated. The width of the broad component of the Balmer line is appreciably reduced when CR acceleration is turned on, but the results are not sensitive to the amount of TH. On the other hand, as one can notice in the right panel, the narrow Balmer line becomes broader when CR acceleration is efficient; the effect is more pronounced when non-negligible TH is taken into account. Hence, the distribution function of neutrals becomes broader mainly because of the scattering with a warmer ion distribution in the far precursor (i.e., where TH is more effective), rather than because of the NRF, which operates only within a few CE interaction lengths from the subshock.

In Fig. 8 we show the FWHM of the broad Balmer line as a function of the CR acceleration efficiency  $\epsilon_{CR}$  (left panel) and as a function of the shock velocity  $V_{sh}$  (right panel). In the left panel, the two curves illustrate the change in the FWHM obtained by assuming full or partial temperature equilibration between electrons and protons in the downstream ( $\beta_{down} = 1$  and  $\beta_{down} = 0.01$ , with  $\beta_{down}$  the ratio between electron and ion temperature), for  $V_{sh} = 4,000 \text{ km s}^{-1}$  and  $n_0 = 0.1 \text{ cm}^{-3}$ .

The FWHM decreases when the energy density in CRs increases, confirming the very important notion that the width of the broad Balmer line can be used to measure the CR acceleration efficiency. There is however an important caveat to keep in mind: for a given value of  $\epsilon_{CR}$ , the uncertainty in the electron-proton equilibration translates into a spread in the FWHM of  $\sim 500 \text{ km s}^{-1}$ . Hence, unequivocal evidence of efficient CR acceleration is likely to be achievable only if the measured FWHM is below the line corresponding to full equilibration. Even in this case the value of  $\epsilon_{CR}$  is likely to be quite uncertain because of the unknown level of equilibration. For instance, if a FWHM of  $2,200 \text{ km s}^{-1}$  were measured for a shock with  $V_{sh} = 4,000 \text{ km s}^{-1}$ , the result could be interpreted as a measurement of

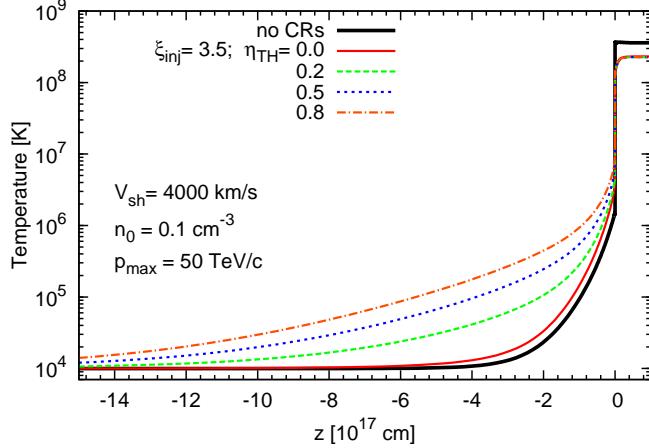


FIG. 6.— Temperature of the plasma in the shock region for  $V_{\text{sh}} = 4,000 \text{ km s}^{-1}$ ,  $n_0 = 0.1 \text{ cm}^{-3}$ ,  $h_N = 0.5$  and  $B_0 = 10 \mu\text{G}$ . When CR acceleration is turned on, we assume  $\xi_{\text{inj}} = 3.5$  and  $p_{\text{max}} = 50 \text{ TeV/c}$ . The thick solid (black) line refers to the case with no CRs. The thin (red) solid line refers to the case with CRs but no TH. Different values of  $\eta_{\text{TH}}$  are as indicated.

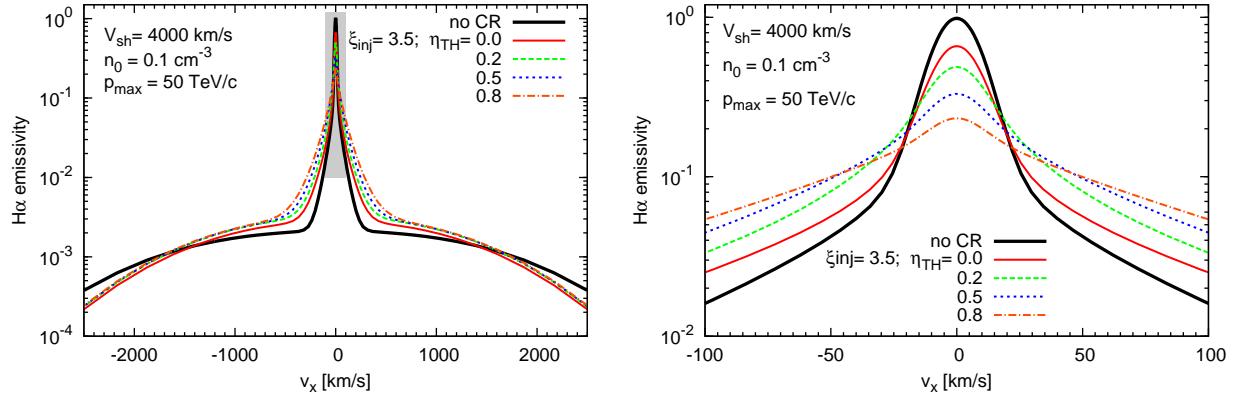


FIG. 7.— Volume-integrated line profile of Balmer emission for the same cases shown in Fig. 6. The right panel shows a zoom on the narrow line region, i.e. the shadow box in the left panel. The thick solid line shows the case without CRs, while different thin lines are calculated at fixed  $\xi_{\text{inj}} = 3.5$  for different values of the TH efficiency.

$\epsilon_{\text{CR}} \sim 0.25$  in the case of full equilibration. In the more general case of partial equilibration, the above value of  $\epsilon_{\text{CR}}$  can be taken as a lower limit. On the other hand, for the same shock velocity, if one measured a FWHM of  $2,600 \text{ km s}^{-1}$ , it would be hardly possible to say anything at all about the efficiency of CR acceleration, without knowing the level of electron-proton thermalization. In order to have an independent estimate of  $\beta$ , it would be very helpful to pair optical observations with the emission of SNR shocks in the (thermal) X-rays as well.

In the right panel of Fig. 8 we fix the injection parameter  $\xi_{\text{inj}} = 3.5$  and we calculate the FWHM of the broad Balmer line as a function of the shock velocity. The two thin (red) lines refer to the cases with no CR acceleration with  $\beta_{\text{down}} = 0.01$  (upper curve) and  $\beta_{\text{down}} = 1$  (upper curve). The dashed ( $\eta_{\text{TH}} = 0$ ) and dash-dotted lines ( $\eta_{\text{TH}} = 1$ ) refer to the cases with CR acceleration (efficiencies in parentheses), again for  $\beta_{\text{down}} = 0.01$  (upper curves) and  $\beta_{\text{down}} = 1$  (lower curves). From the plot it is clear that the TH does not affect appreciably the width of the broad line. As pointed out above, if the electron-proton thermalization is inefficient, it may be difficult to draw any strong conclusion about the CR acceleration efficiency unless additional information is available to constrain the value of  $\beta_{\text{down}}$ . Nevertheless, measuring a FWHM of the broad line below the solid thick line allows one to put a lower limit on the CR acceleration efficiency.

Additional discrimination power might derive from a combined analysis of the broad and narrow Balmer line components. The FWHM of the narrow Balmer line as a function of the maximum momentum  $p_{\text{max}}$  is plotted in Fig. 9 for  $\eta_{\text{TH}} = 0.2$  (top left panel),  $\eta_{\text{TH}} = 0.5$  (top right panel) and  $\eta_{\text{TH}} = 0.8$  (bottom panel). The three curves in each panel are obtained for  $\xi_{\text{inj}} = 3.5, 3.7, 3.8$ . The shock velocity and total density are set to  $V_{\text{sh}} = 4,000 \text{ km s}^{-1}$  and  $n_0 = 0.1 \text{ cm}^{-3}$ , which approximately correspond to  $\epsilon_{\text{CR}} = 0.4, 0.2$  and  $0.1$  (weakly dependent upon  $\eta_{\text{TH}}$ ) for the three values of  $\xi_{\text{inj}}$ . The maximum momentum,  $p_{\text{max}}$  determines the spatial extent of the CR-induced precursor. Larger values of  $p_{\text{max}}$  imply that there is more time (space) for depositing heat in the upstream, and the width of the narrow

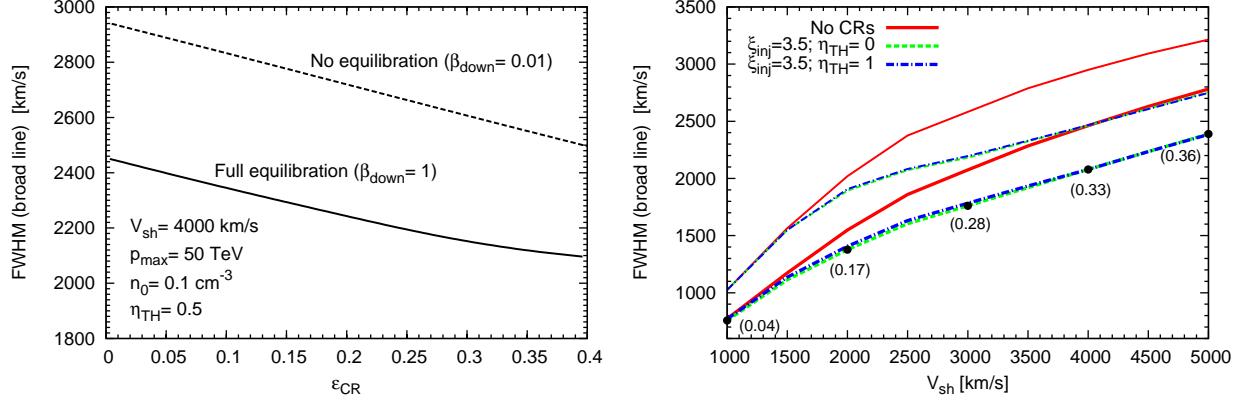


FIG. 8.— FWHM of the broad Balmer line as a function of the CR acceleration efficiency  $\epsilon_{\text{CR}}$  (left panel) and as a function of the shock velocity  $V_{\text{sh}}$  (right panel). The upper set of curves (thin lines) correspond to the case  $\beta_{\text{down}} = 0.01$ , while the lower curves (thick lines) correspond to  $\beta_{\text{down}} = 1$ . In the right panel, the values near the points refer to the associated CR acceleration efficiency.

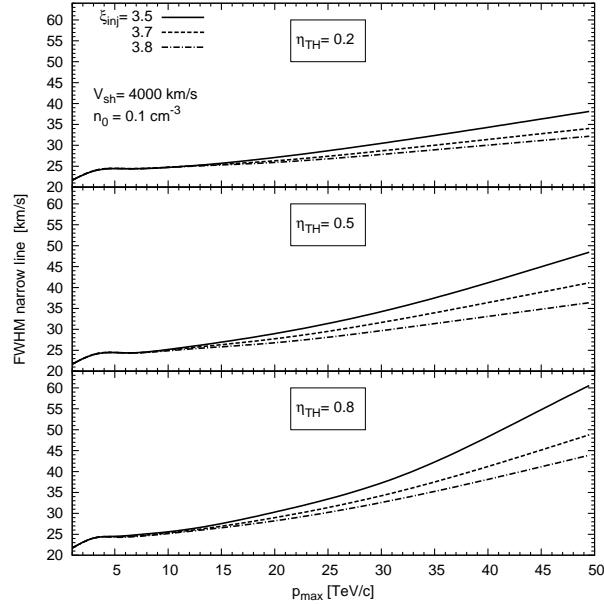


FIG. 9.— FWHM of the narrow line as a function of the maximum momentum of accelerated protons. The three panels refer to  $\eta_{\text{TH}} = 0.2$ , 0.5 and 0.8. Each panel shows three lines calculated assuming different injection parameter,  $\xi_{\text{inj}} = 3.5, 3.7$  and 3.8 which approximately correspond to  $\epsilon_{\text{CR}} = 0.4, 0.2$  and 0.1.

Balmer line broadens correspondingly. The effect becomes more pronounced for larger values of the parameter  $\eta_{\text{TH}}$ . It is worth noticing that the behavior of the curves in Fig. 9 is the same in all cases and reflects the slight broadening of the narrow Balmer line as due to the NRF. In fact, one can estimate the momentum at which the diffusion length equals the size of the precursor induced by the NRF, which is of order  $\lambda_{\text{NRF}} \sim 10^{16}$  cm for the parameters used here. By imposing  $D(p_{\text{max}})/V_{\text{sh}} = p_{\text{max}}c^2/3e\delta B = \lambda_{\text{NRF}}$  one gets  $p_{\text{max}} \sim 5 - 10$  TeV/c (from our calculations the amplified magnetic field close to the subshock is  $\delta B \sim 50\mu\text{G}$ ). This interpretation of the modest increase in the FWHM for low values of  $p_{\text{max}}$  is further confirmed by the fact that the width is very weakly dependent on both  $\xi_{\text{inj}}$  and  $\eta_{\text{TH}}$  for  $p_{\text{max}} \lesssim 10$  TeV/c.

For completeness, Figure 10 shows the width of the intermediate component of the Balmer line as a function of the CR acceleration efficiency  $\epsilon_{\text{CR}}$ , for  $V_{\text{sh}} = 4,000 \text{ km s}^{-1}$ ,  $\eta_{\text{TH}} = 0.5$  and neutral fraction  $h_N = 0.5$ . The two curves refer again to  $\beta_{\text{down}} = 0.01$  (upper curve) and  $\beta_{\text{down}} = 1$  (lower curve). In the upstream, energy is assumed to be deposited only in the protons, which therefore become warmer than the electrons. We assume that they do not equilibrate very effectively since the Coulomb collision timescale is typically longer than the advection one.

The reason why the width of the intermediate Balmer line depends on the CR acceleration efficiency is that when  $\epsilon_{\text{CR}}$  increases, the amount of TH close to the subshock increases as well, leading to a broader distribution of neutrals in the region affected by the NRF. The dependence of the width of the intermediate line on  $\beta_{\text{down}}$  lies in the fact that the NRF carries upstream the information about the ion temperature behind the shock. We further notice that the

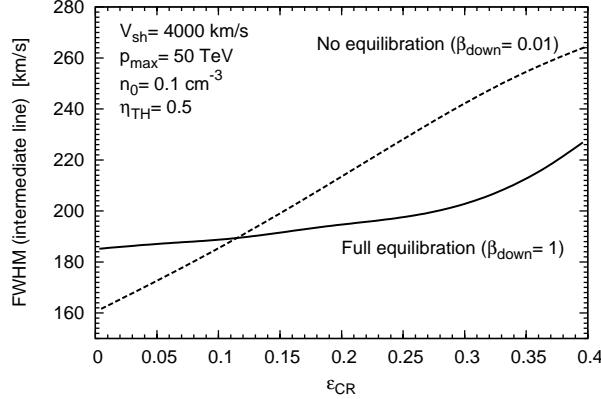


FIG. 10.— FWHM of the intermediate Balmer line as a function of the CR acceleration efficiency,  $\epsilon_{\text{CR}}$ , for a shock speed  $V_{\text{sh}} = 4,000 \text{ km s}^{-1}$   $\eta_{\text{TH}} = 0.5$  and  $p_{\text{max}} = 50 \text{ TeV}$ . The solid and dashed lines refer to the cases  $\beta_{\text{down}} = 1$  and  $\beta_{\text{down}} = 0.01$  respectively.

width of the intermediate component may also depend on the electron-ion equilibration upstream,  $\beta_{\text{up}}$ . In Figure 10 we assume no equilibration upstream, namely while ions are heated by NRF and TH, electrons always have  $T = 10^4 \text{ K}$ . Hence one has to be careful in using the intermediate line to infer shock properties, because this line is a function of several parameters.

#### 4. SUMMARY AND CONCLUSIONS

In this paper we have presented the first formulation of the non-linear theory of diffusive shock acceleration at SNR shocks propagating in a partially ionized medium. At the same time, we have put forward the first fully self-consistent calculation of how the shape of Balmer  $H_{\alpha}$  line is affected by the presence of accelerated particles. This development is crucial if one wants to use the width of the various components of the Balmer line to infer information on the efficiency of CR acceleration in SNRs.

The three most important physical points discussed here in a quantitative way are the following: 1) neutrals produced in CE events with hot ions downstream can return upstream and deposit energy and momentum there. This phenomenon, that we refer to as the neutral return flux (NRF), has an important dynamical effect on the structure of collisionless shocks as well as on the spectrum of accelerated particles (see also Paper I). 2) The structure of a collisionless shock is affected in a non-trivial way by the combined action of neutrals and CRs. 3) Also the shape of the Balmer  $H_{\alpha}$  line reflects the effects of the two phenomena mentioned above.

The NRF was not predicted in previous calculations of the structure of SNR shocks, mainly because of their intrinsic limitation to fluid approaches or restriction to the upstream plasma alone. Both CRs and the NRF produce a precursor upstream of the shock: the CR-induced precursor develops on a scale of the diffusion length of particles at the maximum energy  $\sim D(p_{\text{max}})/V_{\text{sh}}$ , while the neutral-induced precursor develops on a scale of a few CE interaction lengths. Both precursors cause heating of the ion plasma upstream. The heating associated to the NRF is mediated by CE and ionization upstream. The CR-induced heating is most effective only in the presence of turbulent (non adiabatic) heating.

For typical values of the parameters relevant for SNRs, the spatial extent of the neutral-induced precursor is such as to overlap with the diffusion range of particles with  $p \lesssim 5 - 10 \text{ TeV}/c$ . In this energy region the spectrum of accelerated particles is steepened by the dynamical effect of the NRF, as already found in Paper I, in the context of test-particle calculations. The calculations presented here are a generalization of previous work on NLDSA (Amato & Blasi 2005, 2006; Caprioli et al. 2009) and on the dynamical effects of neutrals on the structure of a collisionless shocks (Blasi et al. 2012). The technique adopted for the solution of the non-linear equations that describe the whole problem is basically an iterative method in two stages: in a first stage, we calculated iteratively the distribution of neutrals and ions (see Sections 2.1, 2.4 and Paper I). The distribution function at a generic location is not Maxwellian therefore the calculation of the neutral distribution is kinetic, while ions (protons and electrons) are treated as fluids. In a second stage, the updated profile of neutrals and ions is used to calculate the spatial distribution and spectrum of accelerated particles and waves (Sections 2.2 and 2.3). The procedure is repeated until convergence is reached.

On average, as expected, the presence of neutrals leads to less concave and steeper spectra of accelerated particles than in a completely ionized medium. These effects are quite substantial for shocks with  $V_{\text{sh}} \lesssim 3,000 \text{ km s}^{-1}$ . For faster shocks, the NRF becomes dynamically less important because neutrals are ionized downstream before they suffer a CE reaction. In these cases the shock modification is mainly induced by CRs.

Besides the general interest in a theory of NLDSA in the presence of neutrals, we think that the present work is especially important in terms of providing an essential tool to interpret the shape of the  $H_{\alpha}$  line and translate it into quantitative information on the presence of CRs in SNR shocks. The widths of the narrow and broad component of the Balmer line in SNR shocks have long been known to bear information on the temperature of the ions upstream and downstream of the shock respectively (Chevalier & Raymond 1978; van Adelsberg et al. 1978). Both quantities are

sensitive, in turn, to the CR acceleration efficiency: if CRs are efficiently accelerated, the plasma upstream is heated by the CR-induced precursor and the downstream plasma is heated less than in the absence of CRs because part of the ram pressure is channelled into CR acceleration. In terms of Balmer line, one expects the narrow component to be somewhat broader and the broad component to be somewhat narrower whenever CR acceleration is efficient. A phenomenological approach to the problem was recently put forward by Raymond et al. (2011), where however the CR spectrum and spatial distribution are both assumed rather than calculated. As described in the present paper, the CR properties are profoundly affected by the presence of neutrals. Moreover Raymond et al. (2011) do not treat the downstream plasma, therefore there is no NRF in their calculation.

In Fig. 6 we can see the temperature of the plasma calculated self-consistently. Both the heating in the precursor (for different values of the efficiency of TH) and the reduced temperature in the downstream (as a result of CR acceleration) are clearly visible.

These effects reflect in the shape of the Balmer line, which we calculated by using the formalism introduced in Paper II. The width of the narrow Balmer line is found to depend rather sensibly upon the level of TH and the value of maximum momentum  $p_{\max}$ . The dependence on  $p_{\max}$  can be understood, because increasing  $p_{\max}$  leads to larger CR precursors and therefore neutrals incoming from upstream infinity have a longer time to experience CE and adapt to the ion temperature. Larger TH reflects in an increased width of the narrow Balmer line.

The level of TH does not affect appreciably the plasma temperature downstream, therefore the width of the broad Balmer line mainly reflects the CR acceleration efficiency. However, the broad line is sensitive to the level of equilibration between protons and electrons downstream: if electrons and protons reach thermal equilibrium in a time short compared with CE (clearly this cannot happen through collisions) then the ion temperature is lower, thereby mimicking the presence of CRs. The estimate of the CR acceleration efficiency  $\epsilon_{\text{CR}}$  might be less dependent upon this ambiguity if the width of the broad Balmer line is quite small. These cases however, suggesting a very high acceleration efficiency, are very difficult to obtain in the presence of TH and magnetic reconnection on the shock (Caprioli et al. 2008, 2009). Moreover these cases would correspond to very concave spectra of accelerated particles and of the associated gamma rays, that so far have never been observed. The combined observation of the narrow and broad Balmer line might also help breaking the degeneracy between CR acceleration efficiency and electron-proton equilibration. By the same token, observation of an X-ray spectrum of thermal origin may independently provide a measurement of the electron temperature.

Finally, we stress that the existence of the NRF leads to the formation of an intermediate component of the Balmer line, with a width that depends on the temperature of protons downstream, on the electron-proton equilibration and on the level of TH. The practical possibility to infer  $\epsilon_{\text{CR}}$  from data on individual SNR will need to be assessed on the case-by-case basis.

#### ACKNOWLEDGMENTS

This work was partially funded through grant ASI-INAF I/088/06/0 and PRIN INAF 2010. The research work of D.C. was partially supported by NSF grant AST-0807381.

#### REFERENCES

van Adelsberg, M., Heng, K., McCray, R., and Raymond, J. C. 2008, *ApJ*, 689, 1089  
 Amato, E. & Blasi, P., 2005, *MNRAS*, 364, L76  
 Amato, E. & Blasi, P., 2006, *MNRAS*, 371, 1251  
 Bell, A.R., 1978, *MNRAS*, 182, 147  
 Berezhko, E.G., and Ellison, D.C., 1999, *ApJ*, 526, 385  
 Berezhko, E.G. and Völk, H.J., 2007, *ApJ*, 661, L175  
 Blasi, P., 2002, *Astropart. Phys.*, 16, 429  
 Blasi, P., and Amato, E., 2012, *JCAP*, 01, 10  
 Blasi, P., and Amato, E., 2012, *JCAP*, 01, 11  
 Blasi, P., Amato, A., and Caprioli, D., 2007, *MNRAS*, 375, 1471  
 Blasi, P., Gabici, S., Vannoni, G., 2005, *MNRAS*, 361, 907  
 Blasi, P., Morlino G., Bandiera R., Amato, E. & Caprioli, D. 2012, *ApJ*, 755, 121 [Paper I]  
 Caprioli, D., 2012, *JCAP*, 7, 38  
 Caprioli, D., Blasi, P., Amato, E. & Vietri, M. 2008, *ApJ*, 679, L139  
 Caprioli, D., Blasi, P., Amato, E. & Vietri, M. 2009, *MNRAS*, 395, 895  
 Caprioli, D., Blasi, P. & Amato, 2010, *Astropart. Phys.*, 33, 307  
 Chevalier, R. A., and Raymond, J. C. 1978, *ApJ*, 225, L27  
 Decourchelle, A., Ellison, D. C. & Ballet, J. 2000, *ApJ*, 543, L57  
 Helder, E. A., Kosenko, D., Vink, J. 2010, *ApJ*, 719, L140  
 Helder, E. et al. 2009, *Science*, 325, 719  
 Heng, K., 2009, *PASA*, 27, 23  
 Malkov, M.A. and Drury, L. O'C., 2001, *Rep. Prog. Phys.*, 64, 429  
 Morlino G., Bandiera R., Blasi, P. & Amato, E. 2012, 760, 137 [Paper II]  
 Morlino, G., Amato, E., Blasi, P., Caprioli, D. 2010, proceedings of the 12th ICATPP Conference  
 Ptuskin, V.S., Jones, F.C., Seo, E.S., and Sina, R., 2006, *Adv. in Sp. Res.*, 37, 1909  
 Raymond, J.C., Vink, J., Helder, E.A. and de Laat, A, 2011, *ApJ*, 731, L14  
 Sollerma, J., Ghavamian, P., Lundqvist, P. & Smith, R. C. 2003, *A&A*, 407, 249  
 Vink, J., 2012, *Astr. Astrophys. Rev.*, 20, 49